

56 50

PCT09

RAW SEQUENCE LISTING

DATE: 10/11/2001

PATENT APPLICATION: US/09/646,925

TIME: 10:28:36

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\10112001\I646925.raw

3 <110> APPLICANT: CHATFIELD, STEVEN N.
 5 <120> TITLE OF INVENTION: ATTENUATED BACTERIA USEFUL IN VACCINES
 7 <130> FILE REFERENCE: 117-320
 9 <140> CURRENT APPLICATION NUMBER: US 09/646,925
 C--> 10 <141> CURRENT FILING DATE: 2001-01-31 8✓
 12 <150> PRIOR APPLICATION NUMBER: PCT/GB99/00935
 13 <151> PRIOR FILING DATE: 1999-03-25
 15 <150> PRIOR APPLICATION NUMBER: GB 9806449.6
 16 <151> PRIOR FILING DATE: 1998-03-25
 18 <160> NUMBER OF SEQ ID NOS: 28
 20 <170> SOFTWARE: PatentIn Ver. 2.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 1690
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Escherichia coli
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (492)..(1562)
 31 <400> SEQUENCE: 1
 32 gtcgacgcgg tggatatctc tccagacgcg ctggcggttg ctgaacagaa catcgaagaa 60
 34 cacggtctga tccacaacgt cattccgatt cgttccgcatc tgttccgcga cttgccgaaa 120
 36 gtgcagtacg acctgattgt cactaaccgc cgtatgtcg atgcgaagat atgtccgacc 180
 38 tgccaaacaa taccgccacg agccggaact gggcctggca tctggcactg acggcctgaa 240
 40 actgacgcgt cgcattctcg gtaacgcggc agattacctt gctgatgatg gcgtgttgat 300
 42 ttgtgaagtc ggcaacagca tggtagatct tatggaacaa tatccgatg ttccgttcac 360
 44 ctggctggag ttgataacg gcggcgatgg tgtgtttatg ctaccaaag agcagcttat 420
 46 tgccgcacga gaacatttcg cgatttataa agattaagta aacacgcaaa cacaacaata 480
 48 acggagccgt g atg gct gga aac aca att gga caa ctc ttt cgc gta acc 530
 49 Met Ala Gly Asn Thr Ile Gly Gln Leu Phe Arg Val Thr
 50 1 5 10
 52 acc ttc ggc gaa tcg cac ggg ctg gcg ctc ggc tgc atc gtc gat ggt 578
 53 Thr Phe Gly Glu Ser His Gly Leu Ala Leu Gly Cys Ile Val Asp Gly
 54 15 20 25
 56 gtt ccg cca ggc att ccg ctg acg gaa gcg gac ctg caa cat gac ctc 626
 57 Val Pro Pro Gly Ile Pro Leu Thr Glu Ala Asp Leu Gln His Asp Leu
 58 30 35 40 45
 60 gac cgt cgt cgc cct ggg aca tcg cgc tat acc acc cag cgc cgc gag 674
 61 Asp Arg Arg Arg Pro Gly Thr Ser Arg Tyr Thr Thr Gln Arg Arg Glu
 62 50 55 60
 64 ccg gat cag gtc aaa att ctc tcc ggt gtt ttt gaa ggc gtt act acc 722
 65 Pro Asp Gln Val Lys Ile Leu Ser Gly Val Phe Glu Gly Val Thr Thr
 66 65 70 75
 68 ggc acc agc att ggc ttg ttg atc gaa aac act gac cag cgc tct cag 770
 69 Gly Thr Ser Ile Gly Leu Leu Ile Glu Asn Thr Asp Gln Arg Ser Gln
 70 80 85 90
 72 gat tac agt gcg att aag gac gtt ttc cgt cca ggc cat gcc gat tac 818
 73 Asp Tyr Ser Ala Ile Lys Asp Val Phe Arg Pro Gly His Ala Asp Tyr

ENTERED

RAW SEQUENCE LISTING

DATE: 10/11/2001

PATENT APPLICATION: US/09/646,925

TIME: 10:28:36

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\10112001\I646925.raw

```

74      95      100      105
76 acc tac gaa caa aaa tac ggt ctg cgc gat tat cgc ggc ggt gga cgt 866
77 Thr Tyr Glu Gln Lys Tyr Gly Leu Arg Asp Tyr Arg Gly Gly Gly Arg
78 110      115      120      125
80 tct tcc gcc cgc gaa acc gcc atg cgc gtg gcg gca gga gct att gcc 914
81 Ser Ser Ala Arg Glu Thr Ala Met Arg Val Ala Ala Gly Ala Ile Ala
82      130      135      140
84 aaa aaa tat ctc gcc gag aaa ttt ggt att gaa atc cgt ggc tgc ctg 962
85 Lys Lys Tyr Leu Ala Glu Lys Phe Gly Ile Glu Ile Arg Gly Cys Leu
86      145      150      155
88 acc cag atg ggc gac att ccg ctg gat atc aaa gac tgg tgc cag gtc 1010
89 Thr Gln Met Gly Asp Ile Pro Leu Asp Ile Lys Asp Trp Ser Gln Val
90      160      165      170
92 gag caa aat ccg ttt ttt tgc ccg gac ccc gac aaa atc gac gcg tta 1058
93 Glu Gln Asn Pro Phe Phe Cys Pro Asp Pro Asp Lys Ile Asp Ala Leu
94      175      180      185
96 gac gag ttg atg cgt gcg ctg aaa aaa gag ggc gac tcc atc ggc gct 1106
97 Asp Glu Leu Met Arg Ala Leu Lys Lys Glu Gly Asp Ser Ile Gly Ala
98 190      195      200      205
100 aaa gtc acc gtt gtt gcc agt ggc gtt cct gcc gga ctt ggc gag ccg 1154
101 Lys Val Thr Val Val Ala Ser Gly Val Pro Ala Gly Leu Gly Glu Pro
102      210      215      220
104 gtc ttt gac cgc ctg gat gct gac atc gcc cat gcg ctg atg agc atc 1202
105 Val Phe Asp Arg Leu Asp Ala Asp Ile Ala His Ala Leu Met Ser Ile
106      225      230      235
108 aac gcg gtg aaa ggc gtg gaa att ggc gac ggc ttt gac gtg gtg gcg 1250
109 Asn Ala Val Lys Gly Val Glu Ile Gly Asp Gly Phe Asp Val Val Ala
110      240      245      250
112 ctg cgc ggc agc cag aac cgc gat gaa atc acc aaa gac ggt ttc cag 1298
113 Leu Arg Gly Ser Gln Asn Arg Asp Glu Ile Thr Lys Asp Gly Phe Gln
114      255      260      265
116 agc aac cat gcg ggc ggc att ctc ggc ggt atc agc agc ggg cag caa 1346
117 Ser Asn His Ala Gly Gly Ile Leu Gly Gly Ile Ser Ser Gly Gln Gln
118 270      275      280      285
120 atc att gcc cat atg gcg ctg aaa ccg acc tcc agc att acc gtg ccg 1394
121 Ile Ile Ala His Met Ala Leu Lys Pro Thr Ser Ser Ile Thr Val Pro
122      290      295      300
124 ggt cgt acc att aac cgc ttt ggc gaa gaa gtt gag atg atc acc aaa 1442
125 Gly Arg Thr Ile Asn Arg Phe Gly Glu Glu Val Glu Met Ile Thr Lys
126      305      310      315
128 ggc cgt cac gat ccc tgt gtc ggg atc cgc gca gtg ccg atc gca gaa 1490
129 Gly Arg His Asp Pro Cys Val Gly Ile Arg Ala Val Pro Ile Ala Glu
130      320      325      330
132 gcg aat gct ggc gat cgt ttt aat gga tca cct gtt acg gca acg ggc 1538
133 Ala Asn Ala Gly Asp Arg Phe Asn Gly Ser Pro Val Thr Ala Thr Gly
134      335      340      345
136 gca aaa tgc cga tgt gaa gac tga tattccacgc tggtaaaaaa tgaataaaac 1592
137 Ala Lys Cys Arg Cys Glu Asp
138 350      355

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/646,925

DATE: 10/11/2001

TIME: 10:28:36

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\10112001\I646925.raw

```

140 cgcgattgcg ctgctggtctc tgcttgccag tagcgccagc ctggcagcga cgccgtggca 1652
142 aaaaataaacc caacctgtgc cgggtagcgc caaatcga 1690
145 <210> SEQ ID NO: 2
146 <211> LENGTH: 356
147 <212> TYPE: PRT
148 <213> ORGANISM: Escherichia coli
150 <400> SEQUENCE: 2
151 Met Ala Gly Asn Thr Ile Gly Gln Leu Phe Arg Val Thr Thr Phe Gly
152 1 5 10 15
154 Glu Ser His Gly Leu Ala Leu Gly Cys Ile Val Asp Gly Val Pro Pro
155 20 25 30
157 Gly Ile Pro Leu Thr Glu Ala Asp Leu Gln His Asp Leu Asp Arg Arg
158 35 40 45
160 Arg Pro Gly Thr Ser Arg Tyr Thr Thr Gln Arg Arg Glu Pro Asp Gln
161 50 55 60
163 Val Lys Ile Leu Ser Gly Val Phe Glu Gly Val Thr Thr Gly Thr Ser
164 65 70 75 80
166 Ile Gly Leu Leu Ile Glu Asn Thr Asp Gln Arg Ser Gln Asp Tyr Ser
167 85 90 95
169 Ala Ile Lys Asp Val Phe Arg Pro Gly His Ala Asp Tyr Thr Tyr Glu
170 100 105 110
172 Gln Lys Tyr Gly Leu Arg Asp Tyr Arg Gly Gly Gly Arg Ser Ser Ala
173 115 120 125
175 Arg Glu Thr Ala Met Arg Val Ala Ala Gly Ala Ile Ala Lys Lys Tyr
176 130 135 140
178 Leu Ala Glu Lys Phe Gly Ile Glu Ile Arg Gly Cys Leu Thr Gln Met
179 145 150 155 160
181 Gly Asp Ile Pro Leu Asp Ile Lys Asp Trp Ser Gln Val Glu Gln Asn
182 165 170 175
184 Pro Phe Phe Cys Pro Asp Pro Asp Lys Ile Asp Ala Leu Asp Glu Leu
185 180 185 190
187 Met Arg Ala Leu Lys Lys Glu Gly Asp Ser Ile Gly Ala Lys Val Thr
188 195 200 205
190 Val Val Ala Ser Gly Val Pro Ala Gly Leu Gly Glu Pro Val Phe Asp
191 210 215 220
193 Arg Leu Asp Ala Asp Ile Ala His Ala Leu Met Ser Ile Asn Ala Val
194 225 230 235 240
196 Lys Gly Val Glu Ile Gly Asp Gly Phe Asp Val Val Ala Leu Arg Gly
197 245 250 255
199 Ser Gln Asn Arg Asp Glu Ile Thr Lys Asp Gly Phe Gln Ser Asn His
200 260 265 270
202 Ala Gly Gly Ile Leu Gly Gly Ile Ser Ser Gly Gln Gln Ile Ile Ala
203 275 280 285
205 His Met Ala Leu Lys Pro Thr Ser Ser Ile Thr Val Pro Gly Arg Thr
206 290 295 300
208 Ile Asn Arg Phe Gly Glu Glu Val Glu Met Ile Thr Lys Gly Arg His
209 305 310 315 320
211 Asp Pro Cys Val Gly Ile Arg Ala Val Pro Ile Ala Glu Ala Asn Ala
212 325 330 335

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/646,925

DATE: 10/11/2001

TIME: 10:28:36

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\10112001\I646925.raw

```

214 Gly Asp Arg Phe Asn Gly Ser Pro Val Thr Ala Thr Gly Ala Lys Cys
215          340          345          350
217 Arg Cys Glu Asp
218          355
220 <210> SEQ ID NO: 3
221 <211> LENGTH: 1713
222 <212> TYPE: DNA
223 <213> ORGANISM: Escherichia coli
225 <220> FEATURE:
226 <221> NAME/KEY: CDS
227 <222> LOCATION: (491)..(1594)
229 <400> SEQUENCE: 3
230 gttacaagc gttatagttt ttctgtggtg gcacagaata atgaaaagtg tgtaaagaag 60
232 ggtaaaaaaa accgaatgcg aggcattccg ttgaaatagg ggtaaacaga cattcagaaa 120
234 tgaatgacgg taataaataa agttaatgat gatagcggga gttattctag ttgcgagtga 180
236 aggttttgtt ttgacattca gtgctgtcaa ataactaaga ataagttatt gattttaacc 240
238 ttgaattatt attgcttgat gttagggtgct tatttcgcca ttccgcaata atcttaaaaa 300
240 gttcccttgc atttacattt tgaacatct atagcgataa atgaaacatc ttaaaagtgt 360
242 tagtatcata ttctgtgttg attattctgc atttttgggg agaattggact tgccgactga 420
244 ttaatgaggg ttaatcagta tgcagtggca taaaaaagca aataaaggca tataacagag 480
246 ggtaataaac atg aaa gtt aaa gta ctg tcc ctg ctg gtc cca gct ctg 529
247          Met Lys Val Lys Val Leu Ser Leu Leu Val Pro Ala Leu
248          1          5          10
250 ctg gta gca ggc gca gca aac gct gct gaa gtt tac aac aaa gac ggc 577
251 Leu Val Ala Gly Ala Ala Asn Ala Ala Glu Val Tyr Asn Lys Asp Gly
252          15          20          25
254 aac aaa tta gat ctg tac ggt aaa gta gac ggc ctg cac tat ttc tct 625
255 Asn Lys Leu Asp Leu Tyr Gly Lys Val Asp Gly Leu His Tyr Phe Ser
256          30          35          40          45
258 gac aac aaa gat gta gat ggc gac cag acc tac atg cgt ctt ggc ttc 673
259 Asp Asn Lys Asp Val Asp Gly Asp Gln Thr Tyr Met Arg Leu Gly Phe
260          50          55          60
262 aaa ggt gaa act cag gtt act gac cag ctg acc ggt tac ggc cag tgg 721
263 Lys Gly Glu Thr Gln Val Thr Asp Gln Leu Thr Gly Tyr Gly Gln Trp
264          65          70          75
266 gaa tat cag atc cag ggc aac agc gct gaa aac gaa aac aac tcc tgg 769
267 Glu Tyr Gln Ile Gln Gly Asn Ser Ala Glu Asn Glu Asn Asn Ser Trp
268          80          85          90
270 acc cgt gtg gca ttc gca ggt ctg aaa ttc cag gat gtg ggt tct ttc 817
271 Thr Arg Val Ala Phe Ala Gly Leu Lys Phe Gln Asp Val Gly Ser Phe
272          95          100          105
274 gac tac ggt cgt aac tac ggc gtt gtt tat gac gta act tcc tgg acc 865
275 Asp Tyr Gly Arg Asn Tyr Gly Val Val Tyr Asp Val Thr Ser Trp Thr
276          110          115          120          125
278 gac gta ctg cca gaa ttc ggt ggt gac acc tac ggt tct gac aac ttc 913
279 Asp Val Leu Pro Glu Phe Gly Gly Asp Thr Tyr Gly Ser Asp Asn Phe
280          130          135          140
282 atg cag cag cgt ggt aac ggc ttc gcg acc tac cgt aac act gac ttc 961
283 Met Gln Gln Arg Gly Asn Gly Phe Ala Thr Tyr Arg Asn Thr Asp Phe

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/646,925

DATE: 10/11/2001

TIME: 10:28:36

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\10112001\I646925.raw

```

284          145          150          155
286 ttc ggt ctg gtt gac ggc ctg aac ttt gct gtt cag tac cag ggt aaa 1009
287 Phe Gly Leu Val Asp Gly Leu Asn Phe Ala Val Gln Tyr Gln Gly Lys
288          160          165          170
290 aac ggc aac cca tct ggt gaa ggc ttt act agt ggc gta act aac aac 1057
291 Asn Gly Asn Pro Ser Gly Glu Gly Phe Thr Ser Gly Val Thr Asn Asn
292          175          180          185
294 ggt cgt gac gca ctg cgt caa aac ggc gac ggc gtc ggc ggt tct atc 1105
295 Gly Arg Asp Ala Leu Arg Gln Asn Gly Asp Gly Val Gly Gly Ser Ile
296 190          195          200          205
298 act tat gat tac gaa ggt ttc ggt atc ggt ggt gcg atc tcc agc tcc 1153
299 Thr Tyr Asp Tyr Glu Gly Phe Gly Ile Gly Gly Ala Ile Ser Ser Ser
300          210          215          220
302 aaa cgt act gat gct cag aac acc gct gct tac atc ggt aac ggc gac 1201
303 Lys Arg Thr Asp Ala Gln Asn Thr Ala Ala Tyr Ile Gly Asn Gly Asp
304          225          230          235
306 cgt gct gaa acc tac act ggt ggt ctg aaa tac gac gct aac aac atc 1249
307 Arg Ala Glu Thr Tyr Thr Gly Gly Leu Lys Tyr Asp Ala Asn Asn Ile
308          240          245          250
310 tac ctg gct gct cag tac acc cag acc tac aac gca act cgc gta ggt 1297
311 Tyr Leu Ala Ala Gln Tyr Thr Gln Thr Tyr Asn Ala Thr Arg Val Gly
312          255          260          265
314 tcc ctg ggt tgg gcg aac aaa gca cag aac ttc gaa gct gtt gct cag 1345
315 Ser Leu Gly Trp Ala Asn Lys Ala Gln Asn Phe Glu Ala Val Ala Gln
316 270          275          280          285
318 tac cag ttc gac ttc ggt ctg cgt ccg tcc ctg gct tac ctg cag tct 1393
319 Tyr Gln Phe Asp Phe Gly Leu Arg Pro Ser Leu Ala Tyr Leu Gln Ser
320          290          295          300
322 aaa ggt aaa aac ctg ggt cgt ggc tac gac gac gaa gat atc ctg aaa 1441
323 Lys Gly Lys Asn Leu Gly Arg Gly Tyr Asp Asp Glu Asp Ile Leu Lys
324          305          310          315
326 tat gtt gat gtt ggt gct acc tac tac ttc aac aaa aac atg tcc acc 1489
327 Tyr Val Asp Val Gly Ala Thr Tyr Tyr Phe Asn Lys Asn Met Ser Thr
328          320          325          330
330 tac gtt gac tac aaa atc aac ctg ctg gac gac aac cag ttc act cgt 1537
331 Tyr Val Asp Tyr Lys Ile Asn Leu Leu Asp Asp Asn Gln Phe Thr Arg
332          335          340          345
334 gac gct ggc atc aac act gat aac atc gta gct ctg ggt ctg gtt tac 1585
335 Asp Ala Gly Ile Asn Thr Asp Asn Ile Val Ala Leu Gly Leu Val Tyr
336 350          355          360          365
338 cag ttc taa tctcgattga tatcgaacaa gggcctgcgg gccctttttt 1634
339 Gln Phe
341 cattgttttc agcgtacaaa ctcaagttttt tgggtgtactc ttgcgaccgt tcgcatgagg 1694
343 ataatacagc acggaata 1713
346 <210> SEQ ID NO: 4
347 <211> LENGTH: 367
348 <212> TYPE: PRT
349 <213> ORGANISM: Escherichia coli
351 <400> SEQUENCE: 4

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/646,925

DATE: 10/11/2001

TIME: 10:28:37

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\10112001\I646925.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date